

Amendments to the Claims

This listing of claims replaces all prior versions and listings of claims in the application.

Listing of Claims

1-19 (canceled)

20. (currently amended) A method of converting a voice signal as spoken by a source speaker into a converted voice signal the acoustic characteristics thereof resemble those of a target speaker, the method comprising:

a determination step of determining a function for transforming acoustic characteristics of the source speaker into acoustic characteristics close to those of the target speaker on the basis of samples of the voices of the source and target speakers, and

a transformation step of transforming acoustic characteristics of the source speaker voice signal to be converted by applying said transformation function,

wherein said determination step comprises a step of determining a function for conjoint transformation of characteristics of the source speaker relating to the spectral envelope and of characteristics of the source speaker relating to the pitch and in that wherein said transformation step comprises applying said conjoint transformation function,

wherein said step of determining a conjoint transformation function comprises,

a step of analyzing source and target speaker voice samples grouped into frames to obtain for each frame information relating to the spectral envelope and to the pitch,

a step of concatenating information relating to the spectral envelope and information relating to the pitch for each of the source and target speakers,

a step of determining a model representing common acoustic characteristics of source speaker and target speaker voice samples, and

a step of determining said conjoint transformation function from said model and the voice samples, and

wherein said steps of analyzing the source and target speaker voice samples are adapted to produce said information relating to the spectral envelope in the form of cepstral coefficients.

21-22. (canceled)

23. (currently amended) A method according to claim [[21]]20, wherein said analysis steps comprise respectively a step of achieving voice samples models as a summation of an harmonic signal and noise, each achieving step comprising :

a substep of estimating the pitch of the voice samples,

a substep of synchronized analysis of the pitch of each frame, and

a substep of estimating spectral envelope parameters of each frame.

24. (currently amended) A method according to claim [[21]]20, wherein said step of determining a model determines a Gaussian probability density mixture model.

25. (previously presented) A method according to claim 24, wherein said step of determining a model comprises:

a substep of determining a model corresponding to a mixture of Gaussian probability densities, and

a substep of estimating parameters of the mixture of Gaussian probability densities from an estimated maximum likelihood between the acoustic characteristics of the source and target speaker samples and the model.

26. (currently amended) A method according to claim [[21]]20, wherein said step of determining at least one transformation function further includes a step of normalizing the pitch of the frames of source and target speaker samples relative to average values of the pitch of the analyzed source and target speaker samples.

27. (currently amended) A method according to claim [[21]]20, including a step of temporally aligning the acoustic characteristics of the source speaker with the acoustic characteristics of the target speaker, this step being executed before said step of determining a conjoint model.

28. (previously presented) A method according to claim 20, including a step of separating voiced frames and non-voiced frames in the source speaker and target speaker voice samples, said step of determining a conjoint transformation function of the characteristics relating to the spectral envelope and to the pitch being based only on said voiced frames and the method including a step of determining a function for transformation of only the spectral envelope characteristics on the basis only of said non-voiced frames.

29. (previously presented) A method according to claim 20, wherein said step of determining at least one transformation function comprises only said step of determining a conjoint transformation function.

30. (previously presented) A method according to claim 20, wherein said step of determining a conjoint transformation function is achieved on the basis of an estimate of the acoustic characteristics of the target speaker, the achievement of the acoustic characteristics of the source speaker being known.

31. (previously presented) A method according to claim 30, wherein said estimate is the conditional expectation of the acoustic characteristics of the target speaker the achievement of the acoustic characteristics of the source speaker being known.

32. (previously presented) A method according to claim 20, wherein said step of transforming acoustic characteristics of the voice signal to be converted includes:

a step of analyzing said voice signal, grouped into frames, to obtain for each frame information relating to the spectral envelope and to the pitch,

a step of formatting the acoustic information relating to the spectral envelope and to the pitch of the voice signal to be converted, and

a step of transforming the formatted acoustic information of the voice signal to be converted using said conjoint transformation function.

33. (currently amended) A method according to claim 28 ~~in conjunction with claim 13~~, including a step of separating voiced frames and non-voiced frames in the source speaker and target speaker voice samples, said step of determining a conjoint transformation function of the characteristics relating to the spectral envelope and to the pitch being based entirely on said voiced frames and the method including a step of determining a function for transformation of only the spectral envelope characteristics on the basis only of said non-voiced frames, and including a step of separating voiced frames and non-voiced frames in said voice signal to be converted, said transformation step comprising:

a substep of applying said conjoint transformation function only to voiced frames of said signal to be converted, and

a substep of applying said transformation function of the spectral envelope characteristics only to non-voiced frames of said signal to be converted.

34. (previously presented) A method according to claim 32, wherein said step of determining a transformation function comprises only said step of determining a conjoint transformation function, and wherein said transformation step comprises applying said conjoint transformation function to the acoustic characteristics of all the frames of said voice signal to be converted.

35. (previously presented) A method according to claim 20, further including a step of synthesizing a converted voice signal from said transformed acoustic information.

36. (currently amended) A system for converting a voice signal as spoken by a source speaker into a converted voice signal the acoustic characteristics thereof resemble ones of a target speaker, the system comprising:

means for determining at least one function for transforming acoustic characteristics of the source speaker into acoustic characteristics similar to ones of the target speaker on the basis of voice samples as spoken by the source and target speakers,and;

means for transforming acoustic characteristics of the source speaker voice signal to be converted by applying said transformation function $[,]$ ,

wherein said means for determining at least one transformation function comprise a unit for determining a function for conjoint transformation of characteristics of the source speaker relating to the spectral envelope and of characteristics of the source speaker relating to the pitch and ~~in that wherein~~ said transformation means include for applying said conjoint transformation function;

means for analyzing the voice signal to be converted, adapted to produce information relating to the spectral envelope in the form of cepstral coefficients and relating to the pitch of the voice signal to be converted; and

synthesizer means for forming a converted voice signal from at least said spectral envelope and pitch information transformed simultaneously.

37. (canceled)

38. (previously presented) A system according to claim 36, wherein said means for determining an acoustic characteristic transformation function further include a unit for determining at least one transformation function for the spectral envelope of non-voiced frames, said unit for determining the conjoint transformation function being adapted to determine the conjoint transformation function only for voiced frames.